

# Economic Allocation of Reliability Growth Testing Using Weibull Distributions

Mahmoud Awad

Industrial Engineering Department  
American University of Sharjah (AUS)  
Sharjah, United Arab Emirates  
[miawad@aus.edu](mailto:miawad@aus.edu)

Tel: +971 6 515 2984, Fax: +971 6 515 2979

**Abstract**— Reliability growth testing (RGT) has been widely used for assessing the reliability of complex systems in many industries such as automotive, aerospace, and oil and gas industry. The traditional common and practiced approach of RGT is to assess the initial reliability of the system by building and testing few prototypes for a period of time that extends from few months to years. Then, based on the initial reliability, initial testing time, and reliability target; the total testing time is determined using power law based models such as Duane and AMSAA/Crow models. In this paper, a new method is proposed to allocate RGT testing time for both subsystems and system level in order to minimize system failure rate under limited cost and time resources. Unlike existing methods, intensity failure rate is assumed to be dynamic and modeled using Weibull distribution. Modeling using Weibull is more realistic and increases the applicability of the proposed method in real life applications. The proposed method is motivated by real life examples and its effectiveness is demonstrated by real-life examples.

**Keywords**—Reliability Growth Testing; Duane; Repairable systems;

## I. INTRODUCTION

Reliability Growth Testing (RGT) is a valuable tool to measure product reliability improvement either through planned and dedicated testing or through gradual upgrade and improvement of the fielded product [1]. RGT involve planning and execution of test plans, data collection, modeling, fixing, and re-testing to validate reliability requirements of products or systems. The majority of RGT models are either based on Duane empirical graph model, or AMSAA/Crow model which is considered to have higher mathematical fidelity than Duane [2]. Both models are based on the non-homogenous Poisson process (NHPP) assumption which can be modeled using the power law model. Mathematically, this means that the expected number of failures  $E[N(t)]$  at some test time  $t$  can be fitted through a power model such as Weibull as shown below:

$$E[N(t)] = \theta_s t^{\beta_s} \quad (1)$$

where  $\theta_s$  and  $\beta_s$  represent the system Weibull scale and shape parameters, respectively. The failure intensity or failure frequency at time  $t$  can be represented by  $u(t)$  as shown in equation 2 below:

$$u(t) = \beta_s \theta_s t^{\beta_s - 1} \quad (2)$$

Depending on the value of the shape parameter  $\beta_s$ , the failure rate can be decreasing over time (reliability grow) if  $\beta_s < 1.0$  or constant if  $\beta_s = 1.0$  or increasing (reliability decay) if  $\beta_s > 1.0$ . Scale and shape parameters of the intensity function are nearly always estimated using the maximum likelihood (ML) method [2] and can change only if a hardware or software change is imposed on the system which will change the system reliability performance.

For a repairable system following NHPP process, the probability that a system of age  $t$  fails between  $t$  and  $t + \Delta t$  is given by  $u(t)\Delta t$ . Unlike component probability of failure, repairable system probability of failure is not conditioned on no-system failure up to time  $t$ .

To plan an RGT test for a system, initial values for  $\beta_s$  and  $\theta_s$  are assumed to be known at some initial time  $t_I$ . These initial values are obtained based on previous experience of a similar product or after engagement of system testing for some time. Once initial testing time, initial intensity rate, and shape factor  $\beta_s$  is obtained, the final testing time  $t_F$  is calculated as shown below [3]:

$$u(t_F) = u_I \left( \frac{t_F}{t_I} \right)^{\alpha - 1} \quad (3)$$

Hence, the required testing time to demonstrate a certain failure intensity is:

$$t_F = \text{Exp} \left( \frac{\ln(u_F) - \ln(u_I)}{\alpha - 1} + \ln(t_I) \right) \quad (4)$$

The model in equation 3 above uses a growth rate parameter  $\alpha$  which represents the rate of reliability improvement, or failure rate reduction. Mean time between failures (MTBF) can also be used instead of failure intensity as well.

Practitioners tend to follow two different approaches to derive the initial values: physical RGT initial testing and qualitative RGT testing. The first approach is more common where engineering prototypes of systems are built, tested, and initial estimates of failure intensity and shape factors are estimated based on testing time, failure frequency and correction actions made or planned. To shorten the total testing time, accelerated test methods are often used for long life products. However,

using accelerated methods is not practical for all applications because of complexity or lack of understanding of failure mechanisms. During the initial physical RGT testing, it is not uncommon for some organizations to use a correlation factor of two to four between RGT testing and field. The rationale behind such factor is the criticality of testing personnel during RGT and their ability to detect issues or failures sooner than customers in the field. Project managers are usually not in favor of using such approach since it does not help them to allocate resources needed for RGT execution upfront. This issue gets more complicated when building the test units is expensive and time consuming like the case in aerospace, automotive, and oil and gas applications.

In the second approach, i.e. qualitative reliability growth, engineering judgment and previous experience with similar products are used to guess  $u_i$  [4] [5]. This method relies on the expertise of the team involved in the assessment and complexity of the new product developed; hence it could be subjective and risky.

Although the AMSAA/CROW planning has been used extensively in different industries [6] [5] [7] [8], it has two major drawbacks:

1. The model assumes that shape and scale parameters do not change over time. The most common approach for RGT is to deploy testing of systems, collect data, and analyze failures, fix failure modes and re-test again. The model parameters may change over time with more and more corrective actions implemented
2. The model does not account for the variability of the intensity rate  $u(t)$  estimate. Uncertainty may put unwarranted emphasis on a system with great potential for improvement; however this is risky if the failure intensity estimate is not accurate [3]. It might be more beneficial to allocate the resources to a system with lower potential for improvement but with great confidence in the intensity estimate.

Moreover, the 2008 Defense Science Board Task Force raised a major concern regarding the major risks of the initial assumptions [9]. As a result, many researchers use the Bayesian reliability demonstration theorem [10] [11] [12] to come up with more reliable initial assumptions. The Bayesian method considers various prior information such as historical test data, field data, design changes, corrective actions, and field data for similar products and factors this information in the initial assumptions. However, research on information integration in RGT is relatively scarce [13] and can be subjective [14]. Another enabler approach is the failure mode elimination using stress tests [7] where reliability credit is given based on the ability of the design team to “weed out” a potential failure mode based on the understanding of the physics of the failure mode through detection or prevention risk mitigation.

To tackle the second drawback, i.e. variability consideration, Coit [3] proposed a method to allocate subsystem RGT test time in order to minimize failure intensity and variance of

failure intensity estimate of all subsystems with limited resources. His model is summarized below:

$$\min \sum_{i=1}^s \lambda_i \beta_i (\tau_i + \tau_{0,i})^{\beta_i - 1} + Z_\alpha \left[ \sum_{i=1}^s \sigma_{i,u(\tau_i)} \right]^{\frac{1}{2}}$$

$$s. t. \sum_{i=1}^s C_i \tau_i \leq C$$

$$\tau_i \geq 0 \text{ for } i = 1, \dots, s$$

Where

$\lambda_i$  and  $\beta_i$  are shape and scale parameters for subsystem  $i$   
 $Z_\alpha = (1 - \alpha) \times 100\%$  percentile of a standard normal distribution

$s$  = number of subsystems

$\tau_i$  = additional planned test time for subsystem  $i$

$\tau_{0,i}$  = initial test time already accumulated for subsystem  $i$

$C_i$  = cost of testing subsystem  $i$  per unit time

$C$  = available cost budget for RGT testing

The first term in the objective function represents the summation of subsystem intensity estimates and the second term represents the variance of these estimates. Although the above model provides a major improvement over AMSAA/CROW model, it suffers from the following limitations:

1. It assumes that the estimate of failure intensity (at time  $\tau_i$ ) is normally distributed or number of subsystems  $s$  is large enough to be approximated as normal according to the central limit theorem (CLT) [3] [15]. Using a more flexible and generic distribution such as Weibull distribution is more reasonable especially when dealing with mechanical or electro-mechanical systems [16]. The validity of Weibull distribution will be demonstrated using the motivational example in the next section.
2. The model accounts for RGT testing of subsystems only. Although testing at subsystem level only is very attractive due to cost, in practice RGT testing is done at both subsystem and system level since some of the failure modes do not manifest themselves except at system level. As a matter of fact, researchers classified failures into two categories: surfaced failure modes occur during in-house RGT testing, and latent failure modes occur post system installation or post RGT [15].

The purpose of this article is to propose a new method to allocate RGT testing time that addresses the above two limitations. The new method considers both subsystem and system testing time allocation in order to minimize the system intensity failure rate, i.e. maximizing system reliability while using a more generic distribution for intensity function such as Weibull. This work is motivated by the increasing importance of RGT planning which is recognized by many practitioners and researchers [9] [17].

The rest of this article is organized as follows. The next section will present a real-life example that demonstrates the

need for a new method. Section III outlines the proposed methodology based on the Weibull distribution assumption. Section VI provides a numerical real life example to demonstrate the proposed methodology followed by conclusion in the last section.

## II. MOTIVATIONAL EXAMPLES

Data acquisition systems (DAQ) are complex telemetry systems used to capture operational data from complex systems such as power generators, transformers, airplanes and oil and gas tools. A 650 newly developed DAQ systems were tested for a full year as part of RGT plan. In an attempt to understand the variability of failure intensity rate  $u(t)$  as a function of testing time, the authors analyzed the reliability data for this DAQ 650 systems by dividing them into subgroups of 60's and analyzing the failure rate of the subgroups using the NHPP assumption at 500, 1000, and 1500 hours' time points. At each time point, failure and survival records were considered up to that time point while any records after that were censored. Figure 1 summarizes the results of the failure intensity of each subgroup at the three time points along with fitted curves for the mean and variance of the intensity rates. The primary axis was used for all data points except variance for which the secondary vertical axis was used.

As expected, both failure intensity mean and variance dropped with more testing (reliability grow) and discovered failure modes were mitigated. A closer look at the failure intensities at each time point indicates that the best fit distribution is Weibull and not Normal. Figure 2 and Figure 3 show the Weibull and Normal probability plots of the DAQ failure intensity at the three time points respectively. The p-values of normality test is less than 0.1, indicating that the failure intensities at 500, 1000, and 1500 hours are not normal. Table 1 summarizes the shape and scale parameter estimates for the DAQ subsystem. Although the scale parameter of the Weibull distributions dropped with time, the shape parameter did not change significantly ( $0.5 \pm 0.02$ ).

Table 1 Shape and scale parameter estimates for DAQ example

Statistic	Time Point		
	500	1000	1500
Shape parameter $\beta_{u(t)}$	0.49	0.52	0.5
Scale parameter $\theta_{u(t)}$	3.2E-6	1.1E-6	8.5E-7

The DAQ subsystem was tested as a standalone subsystem thoroughly first, then it was tested part of a full system test. It was found that at least 20% of failure modes were discovered only when the DAQ was tested with other subsystems in real life settings. This behavior was also observed in other equipment's RGT such as construction, automobile, and oil and gas equipment's as well. This is understandable since some failure modes are due to subsystem-subsystem interactions which will not be exposed from subsystem testing level only.

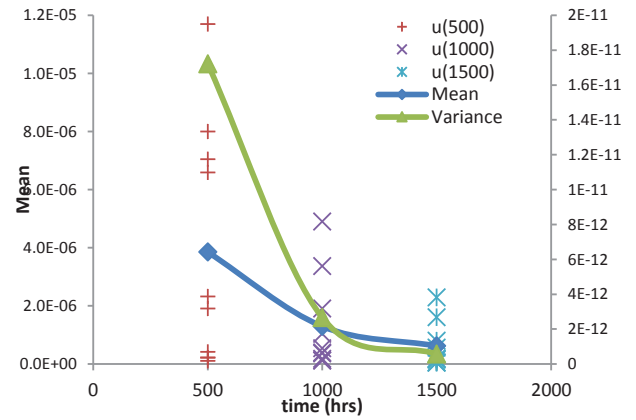


Figure 1 DAQ RGT data analysis

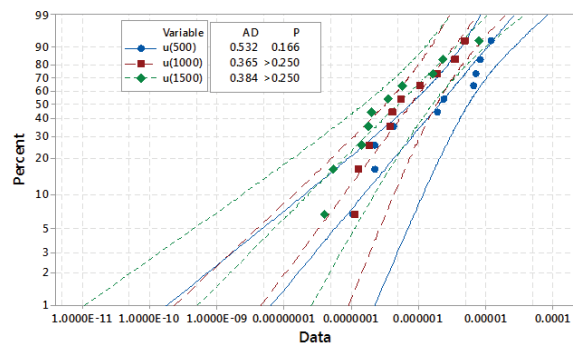


Figure 2 Weibull probability plot for DAQ  $u(t)$

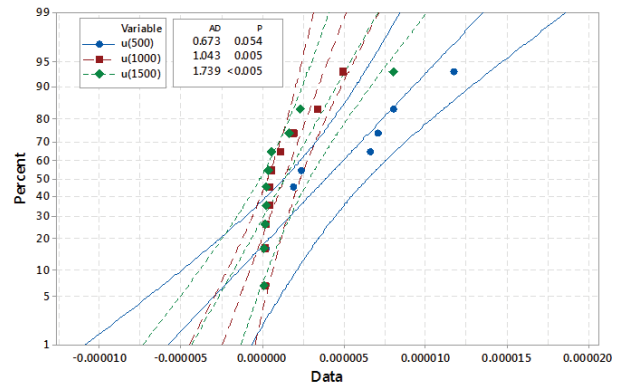


Figure 3 Normal probability plot for DAQ  $u(t)$

## III. PROPOSED METHODOLOGY

The proposed methodology for RGT planning consists of two steps: intensity mean and variance mapping, and testing time

allocation. Each step will be discussed in detail in the next two subsections.

#### A. Failure Intensity Mean and Variance Functions Mapping

The mean of the failure intensity can be described in terms of an exponential function in accordance with Poisson NHPP model shown in equation 1 as follow:

$$\mu_{u(t)} = c_0^\mu e^{c_1^\mu t} \quad (5)$$

where  $\mu_{u(t)}$  is the mean of failure intensity at time  $t$ ,  $c_0^\mu$  and  $c_1^\mu$  are mean coefficients. The mean of failure intensity is expected to drop over time, so the exponential coefficient should be negative. Moreover, a target for failure intensity mean should be set upfront indicating the expected failure intensity at certain testing times. For example, the expected life ( $T$ ) for the DAQ system is 1500 hours; historically the failure intensity for similar products was estimated to be  $1.155 \text{ E-}06$  failures/ hour at 500 hours. The targets were set to lower the intensity by one third at 1000 hours and by half at expected life, i.e. 1500 hours. Figure 4 depicts the DAQ targets and initial historical intensity  $u_0(t)$  along with a fitted exponential equation. As a result,  $c_0^\mu = 2 \times 10^{-6}$  and  $c_1^\mu = -7 \times 10^{-4}$ . Quigley and Walls [18] argued that RGT approach should focus on the identification and realization of faults, which if removed are likely to lead to the improvement in system reliability. RGT planning typically starts with a thorough investigation of existing and anticipated system failure modes and identification of improvement opportunities. Setting failure intensity targets is a natural and reasonable step since areas of improvement and its impact are already defined.

Next the mean targets above will be mapped to Weibull distribution mean equations as follow:

$$u(t_0) = \theta_{u(t_0)} \Gamma\left(1 + \frac{1}{\beta_{u(t_0)}}\right) = u_0 \quad (6)$$

$$u(t_1) = \theta_{u(t_1)} \Gamma\left(1 + \frac{1}{\beta_{u(t_1)}}\right) = u_1 = 0.677u_0 \quad (7)$$

$$u(t_2) = \theta_{u(t_2)} \Gamma\left(1 + \frac{1}{\beta_{u(t_2)}}\right) = u_2 = 0.5u_0 \quad (8)$$

Where  $t_0$ ,  $t_1$ , and  $t_2$  are time points with failure intensities targets are set. In the case of DAQ example,  $t_0$ ,  $t_1$ , and  $t_2$  are 500, 1000, and 1500 hours respectively.  $\theta_{u(t)}$  and  $\beta_{u(t)}$  are the shape and scale parameters of the failure intensity Weibull distributions at the target time points and  $\Gamma(x)$  is the Gamma distribution. Keep in mind that the  $0.677u_0$  and  $0.5u_0$  values are only valid for the DAQ example and can be different depending on the set targets by development team. Without a loss of generality, the shape factor  $\beta_{u(t)}$  will be assumed constant across all time points. This is a fair assumption since the same systems will be used to carry out testing across all data points with additional failure modes show up and

mitigation of old ones. Hence, the overall shape factor will change slightly as shown in DAQ and other examples. Moreover, the gamma function  $\Gamma\left(1 + \frac{1}{\beta_{u(t_2)}}\right)$  is not sensitive to changes relative to the shape parameter.

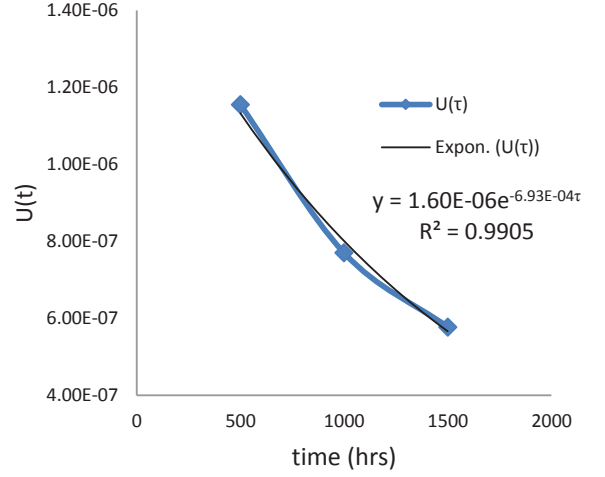


Figure 4 DAQ intensity function targets

For variance of intensity function, we will assume that standard deviation is a function of testing time as shown below:

$$\sigma_{u(t)}(t) = \exp(c_0^\sigma + c_1^\sigma t) \quad (9)$$

Where  $\sigma_{u(t)}$  is the standard deviation of failure intensity at time  $t$ ,  $c_0^\sigma$  and  $c_1^\sigma$  are the standard deviation coefficients.

Similar to intensity mean mapping step above, the standard deviation is mapped to standard deviation of a Weibull distribution at the three target points as shown below:

$$\begin{aligned} \sigma_{u(t_0)} &= \exp(c_0^\sigma + c_1^\sigma t_0) \\ &= \theta_{u(t_0)} \left[ \Gamma\left(1 + \frac{2}{\beta_{u(t_0)}}\right) - \left(\Gamma\left(1 + \frac{1}{\beta_{u(t_0)}}\right)\right)^2 \right] \end{aligned} \quad (10)$$

$$\begin{aligned} \sigma_{u(t_1)} &= \exp(c_0^\sigma + c_1^\sigma t_1) \\ &= \theta_{u(t_1)} \left[ \Gamma\left(1 + \frac{2}{\beta_{u(t_1)}}\right) - \left(\Gamma\left(1 + \frac{1}{\beta_{u(t_1)}}\right)\right)^2 \right] \end{aligned} \quad (11)$$

$$\begin{aligned} \sigma_{u(t_2)} &= \exp(c_0^\sigma + c_1^\sigma t_2) \\ &= \theta_{u(t_2)} \left[ \Gamma\left(1 + \frac{2}{\beta_{u(t_2)}}\right) - \left(\Gamma\left(1 + \frac{1}{\beta_{u(t_2)}}\right)\right)^2 \right] \end{aligned} \quad (12)$$

With the assumption of constant shape parameter, i.e.  $\beta_{u(t_0)} = \beta_{u(t_1)} = \beta_{u(t_2)} = \beta_u$ , equations (6-8) and (10-12) can be used

to solve for the six unknown variables namely:  $\theta_{u(t_0)}, \theta_{u(t_1)}, \theta_{u(t_2)}, \beta_u, c_0^\sigma$ , and  $c_1^\sigma$ . Once these variables are estimated, both mean and variance functions of failure intensity can be formulated. In the case of the DAQ system,  $\theta_{u(t_0)} = 3.68E - 6$ ,  $\theta_{u(t_1)} = 2.12E - 6$ ,  $\theta_{u(t_2)} = 3.84E - 7$ ,  $\beta_u = 0.5$ ,  $c_0^\sigma = -4.64$ , and  $c_1^\sigma = -0.0004$ . One fundamental assumption for the proposed method above to work is that  $\beta_u \leq 1$  which means that the subsystem will become more reliable with testing. If this assumption is violated for any subsystem, this means that the subsystem does not warrant any testing since more testing will make the subsystem less reliable.

### B. Testing Time Allocation

With only one explanatory testing time variable for the location  $\mu_{u(t)}$  and variability  $\sigma_{u(t)}$ , the upper quantile function for the failure intensity function can be expressed as:

$$\log(u_{p(t)}) = \mu_{u(t)} + \Phi_{sev}^{-1}(p)\sigma_{u(t)} \quad (13)$$

where  $\Phi_{sev}^{-1}$  is the  $p^{th}$  percentile of the smallest extreme function. The choice of the smallest extreme function  $\Phi_{sev}$  was based on the fact that  $u(t)$  follows a Weibull distribution. Based on the above, we propose the following problem formulation which is to minimize the lower-bound of the system failure intensity in equation 13. The system is composed of subsystems that work in series and the system intensity failure is the summation of the subsystem intensity functions.

$$\min \sum_{i=1}^s \mu_{i,u(\tau_i)} + \Phi_{sev}^{-1}(p) \left[ \sum_{i=1}^s \sigma_{i,u(\tau_i)}^2 \right]^{\frac{1}{2}}$$

where  $\tau_i = t_{0,i} + D_i t_i + t_{sys}$

$$s.t. \quad C_{sys} t_{sys} + \sum_{i=1}^s C_i \tau_i \leq C$$

$$\sum_{i=1}^s \tau_i \leq T$$

$$t_{sys}, \tau_i \geq 0 \text{ for } i = 1, \dots, s$$

Where,

$s$  = number of subsystems

$\tau_i$  = total testing time of subsystem  $i$

$t_{0,i}$  = initial test time already accumulated for subsystem  $i$

$t_i$  = additional planned testing time of subsystem  $i$

$D_i$  = detection ration of subsystem  $i$  testing  $0 \leq D_i \leq 1$

$t_{sys}$  = planned test time for system

$C_i$  = cost of testing subsystem  $i$  per unit time

$C_{sys}$  = cost of system testing per unit time

$C$  = available cost budget for RGT testing

$T$  = available time budget for RGT testing

The objective of the model is to minimize the system intensity failure which can be obtained as a summation of the intensity failure means of the individual subsystems. It also accounts for intensity variability represented by the square root of the

sum of subsystem variances. The total allocated testing time for each subsystem  $\tau_i$  is made of three components: historical test time  $t_{0,i}$  based on actual accumulated testing or the performance of similar products, additional testing time planned for subsystem  $t_i$  and testing time planned for the whole system  $t_{sys}$ . Since there are certain failure modes resulted from subsystem-subsystem interaction which cannot be detected with individual subsystem testing, a detection ration  $D_i$  ranges between 0 and 1 will be used to give more credit to system testing over subsystem. Since 80% of failure modes were only detected during subsystem testing, a  $D_i = 0.8$  was used for DAQ subsystem. The decision variables for the model is the additional planned subsystem and system testing time, i.e.  $t_i$  and  $t_{sys}$  respectively.

In addition to the non-negativity constrain, two constraints were used to limit total time and cost budget for testing.

## IV. NUMERICAL EXAMPLE

To demonstrate the use of the proposed methodology the following real life example will be used. An oil and gas hydraulic system used for pressure and oil sampling that consists of five subsystems is under development with mission duration of 1500 hours. Due to confidentiality reasons, the original data of each subsystem is altered and presented in Table 2. The total budget dedicated for RGT testing is limited to \$3000,000 and testing time is limited to 18 months (3,168 hours) with system testing cost rate of \$3500/hr. The system testing cost rate is substantially higher than subsystem testing rates for two reasons: system testing is inclusive to all subsystems hardware and requires more staff for operation; and it is done in a real-life setting in a pressure well where subsystem testing is done in lab environment. The last three rows in Table 2 are the results of the mean and variance mapping stage of the proposed method above.

Table 2 Oil & gas hydraulic system RGT data

	Subsystems				
	1	2	3	4 (DAQ)	5
$u_i(t_0)$	1.8E-6	1.5E-7	1.0E-6	1.16E-6	1.7E-6
$t_{0,i}$	500	300	350	500	500
$c_0^\mu$	2.5E-6	1.8E-6	1.23E-6	1.6E-6	2.4E-6
$c_1^\mu$	-6.9E-4	-5.8E-4	-6.04E-4	-6.9E-4	-6.9E-4
$D_i$	0.8	0.8	0.8	0.8	0.8
$C_i$ (\$/hr)	700	900	600	600	800
$\beta_u$	0.6	0.7	0.8	0.5	0.7
$c_0^\sigma$	-5.16	-6.32	-5.17	-4.64	-5.16
$c_1^\sigma$	-3.5E-4	-2.5E-4	-2.7E-4	-3.5E-4	-3.5E-4

Since the objective function is a smooth nonlinear function, the generalized reduced gradient (GRG) solving method [19] with forward derivatives was selected for solving the proposed model. As a matter of fact, the objective function is concave when the decision variables are nonnegative.

Table 3 below show results of baseline model solution in the first row along with 13 more scenarios carried out to investigate sensitivity of solution relative to some input parameters using a 95% upper limit for intensity, i.e.  $p = 0.95$ .

The base line (run 1) indicates that testing on the subsystem level should be done for subsystems one and four for 905 and 1891 hours respectively, followed by system testing for 371 hours. Subsystems one and four had the lowest shape factors  $\beta_u$  and high initial failure intensity  $u_i(t_0)$  which makes them good candidates for improvement. Runs 2-5 had the same parameter values as the base line case with the exception of the detection ration which is varied from 0.6-0.9. It is clear that the lower the detection ration, the less powerful the subsystem testing is, and more reliance on system testing should be done. The same argument can be made for the system cost rate, the lower the system cost rate the higher the system testing allocation is (compare run 3 with 11 and 14). The gradual increase of the shape factor for all subsystems from 0.6-0.9 in runs 6-9 showed a very little effect, if any, on the testing time.

Table 3 Hydraulic system RGT results

run	$\beta_u$	$D$	$C_S$	$t_{sys}$	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$
1	0.6, 0.7, 0.8, 0.5, 0.7	0.8	3500	371	905	0	0	1891	0
2	0.5	0.9	3500	347	1054	0	0	887	880
3	0.5	0.8	3500	348	1070	0	0	890	860
4	0.5	0.7	3500	505	802	0	0	605	546
5	0.5	0.6	3500	1500	0	0	0	0	0
6	0.6	0.8	3500	343	1215	0	11	598	1001
7	0.7	0.8	3500	341	1278	0	92	395	1063
8	0.8	0.8	3500	339	1318	0	144	264	1103
9	0.9	0.8	3500	338	1345	0	178	177	1130
10	0.5	0.9	3000	421	1032	0	0	876	838
11	0.5	0.8	3000	804	443	0	0	271	219
12	0.5	0.7	3000	1000	0	0	0	0	0
13	0.5	0.9	2500	1500	70	0	0	0	0
14	0.5	0.8	2500	1500	0	0	0	0	0

## V. CONCLUSIONS

In this article, a proposed method for estimating RGT testing requirement is presented and demonstrated using a real life numerical example. The method is composed of two stages: intensity mean and variance mapping, followed by testing time allocation using non-linear model with testing time as decision variables. The proposed test method is limited to systems with

continuous time to failure and has potential for reliability growth, i.e. shape factors of less than one.

The new method is based on modeling intensity function at any time using Weibull distribution. It also allows for the allocation of system and subsystem testing at the same time. As a result, it applies to a greater range of systems compared to the previously existing methods.

## REFERENCES

- [1] M. Krasich, "Reliability growth testing, what is the real final result? Accelerated test methodology for reliability growth," in *Reliability and Maintainability Symposium (RAMS 27-30 Jan. 2014)*, Colorado Springs, Co doi: 10.1109/RAMS.2014.6798483, 2014 .
- [2] J. Van Dyck and T. Verdonck, "Precision of power-law NHPP estimates for multiple systems with known failure rate scaling," *Reliability Engineering and System Safety*, vol. 126, pp. 143-152, 2014.
- [3] D. Coit, "Economic allocation of test times for subsystem-level reliability growth testing," *IIE Transactions*, vol. 30, pp. 1143-1151, 1998.
- [4] R. M. Sullivan, *Statistics: Informed Decisions Using Data*, Boston: Prentice Hall, Boston, pp. 7-11., 2011.
- [5] D. Marcus, "Reliability Growth Plan for Orion," Lockheed Martin Space Systems Company Technical Report, 2010.
- [6] M. Krasich, "Power low model, correct application in reliability growth do the cumulative times indeed always add up?," in *RAMS*, Colorado Springs, 2014.
- [7] T. Tekcan, G. Kahramanoglu and M. Gunduzalp, "Determining Reliability by Failure Rate Estimation via a New Parameter," in *Reliability and Maintainability Symposium (RAMS)*, Reno, 2012.
- [8] O. Yadav, N. Singh and P. Goel, "Reliability demonstration test planning: a three dimensional consideration," *Reliability Engineering and System Safety*, vol. 91, no. 8, pp. 882-93, 2006.
- [9] L. Crow, "Demonstrating reliability growth requirements with confidence," in *Reliability and Maintainability Symposium (RAMS)*, Colorado Springs, 2014.
- [10] L. Walls, "Building prior distributions to support Bayesian reliability growth modelling using expert judgement," *Reliability engineering & system safety*, vol. 74, no. 2, pp. 117-128, 2001.
- [11] H. F. Martz, R. A. Waller and E. T. Fickas, " Bayesian Reliability Analysis of Series Systems of Binomial Subsystems and Components," *Technometrics*, vol. 30, no. 2, pp. 143-155, 1988.
- [12] F. Coolen and P. Coolen-Schrijner, *Bayesian Reliability Demonstration*, Wiley Encyclopedia of Statistics in Quality and Reliability, 2008.
- [13] G. Jin and D. Matthews, "Reliability Demonstration for Long-Life Products Based on Degredation Testing and a

- Wiener process Model," *IEEE Transactions on Reliability*, vol. 63, no. 3, pp. 781-797, 2014.
- [14] Peng, W., H.-Z. Huang, Y. Li, M. Zuo and M. Xie, "Life cycle reliability assessment of new products—A Bayesian model updating approach," *Reliability Engineering and System Safety*, vol. 112, pp. 109-119, 2013.
- [15] J. Tongdan and Y. Ying, "A Multiphase Decision Model for System Reliability Growth With Latent Failures," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 43, no. 4, pp. 959-965, 2013.
- [16] B. Mengmeng, S. Jian and W. H. X. Shaoping, "Reliability growth test based on mission reliability for hydraulic actuation systems," in *IEEE 10th International Conference on Industrial Informatics*, 10.1109/INDIN.2012.6301201, 2012.
- [17] M. Krasich, "Power Law Model, Correct Application in Reliability Growth, Do the Cumulative Times Indeed Always Add Up?," in *Reliability and Maintainability Symposium (RAMS)*, Colorado Springs, 2014.
- [18] J. Quigley and L. Walls, "Measuring the effectiveness of reliability growth testing," *Quality and Reliability Engineering International*, vol. 15, no. 2, pp. 87-93, 1999.
- [19] L. Lasdon, A. Warren, A. Jain and M. Ratner, "Design and testing of a generalized reduced gradient code of nonlinear programming," *ACM Trans. Math Softw.*, vol. 1, no. 4, pp. 33-50, 1978.
- [20] W. Nelson, "Weibull analysis of reliability data with few or no failures," *Journal of Quality Technology*, vol. 17, no. 3, pp. 140-146, 1985.
- [21] I. 61710, "Power law model-goodness of fit tests and estimation methods," IEC standards, 2013.